DOCUMENT RESUME

ED 036 295 JC 700 042

AUTHOR FEDRICK, ROBERT JOHN
TITLE GENERALIZED FILE MANAGEMENT SYSTEMS: THEIR IMPLICATION FOR A CALIFORNIA JUNIOR COLLEGE DATA BASE SYSTEM.
PUB DATE 69
NOTE 41P.

EDRS PRICE EDRS PRICE MF-$0.25 HC-$2.15
DESCRIPTORS *DATA BASES, *DATA PROCESSING, *INFORMATION SYSTEMS, *JUNIOR COLLEGES, *MANAGEMENT SYSTEMS

ABSTRACT

CRITERIA TO USE IN EVALUATING DATA PROCESSING EFFICIENCY, FACTORS OF FILE AND RECORD DEFINITIONS, CONVENIENCE OF USE FOR NON-PROGRAMMERS, REPORT GENERATING CAPABILITIES, AND CUSTOMER SUPPORT FOR GENERALIZED FILE MANAGEMENT SYSTEMS FOR USE BY THE CALIFORNIA JUNIOR COLLEGES ARE INDICATED BY THE AUTHOR. THE PURCHASE OF SUCH A SYSTEM AT THE STATE LEVEL IS SUGGESTED. THE SEQUENCE OF ACTIVITIES LEADING TO A USABLE MANAGEMENT INFORMATION SYSTEM WOULD BE (1) THE CREATION OF A COMMON JUNIOR COLLEGE DATA BASE, (2) THE CREATION OF A CENTRAL INFORMATION NETWORK, (3) CRITERIA SPECIFIED FOR A GENERALIZED FILE MANAGEMENT SYSTEM, AND (4) AVAILABILITY OF THIS GENERALIZED SYSTEM TO ACCESS, RETRIEVE, AND UPDATE THE DATA BASE ITSELF. (MC)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED036295

UNIVERSITY OF CALIFORNIA

Los Angeles

Generalized File Management Systems: Their Implication For A California Junior College Data Base System

A paper submitted in partial satisfaction of the
requirements for an Information Science Minor

by

Robert John Fedrick

Committee in charge:

Professor Robert M. Hayes, Chairman

1969

JC 700 042

UNIVERSITY OF CALIF.
LOS ANGELES

FEB 13 1970

ERIC CLEARINGHOUSE
FOR JUNIOR COLLEGES

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1

II. NEED FOR A COMMON DATA BASE IN CALIFORNIA JUNIOR COLLEGES 3

III. AN APPROACH TO DATA BASE DESIGN: MANAGEMENT INFORMATION OR INFORMATION MANAGEMENT? . . . 6

IV. GENERALIZED FILE MANAGEMENT SYSTEMS SUPPORTED BY COMPUTER MANUFACTURERS 13

Information Management System (IMS) 13
Generalized Information System (GIS) 17
Information Oriented Language (INFOL) . . . 21
Business EDP Systems Technique (BEST) . . . 23
Integrated Data Store (IDS) 24
MANAGE 25

V. CONCLUSIONS AND RECOMMENDATIONS 27

BIBLIOGRAPHY 38

I. INTRODUCTION

Robert Head, in a recent article, made the amusing statement that "if we could imagine a present day Mark Twain somehow becoming envolved with computer technology, he might be expected to remark that "a management information system is rather like the weather: everybody discusses it, but nobody does much about it" (8:4).

Junior college administrators have been aware of the possibility of management information systems and have shown a great deal of interest in building common data banks to make operational, information systems of all types. Although the topic is widely discussed, Grant in Education and Data Processing reported that as of July 1967, there was no school in the United States known to have operational a total information system (13:36).

Only recently have schools done much about making information systems a reality. A federal program called Projects to Advance Creativity in Education (PACE), last year approved 14.6 million dollars in projects involving data processing systems (13:37). In addition, the National Science Foundation has established a new Office of Computing Activities to encourage the use of computers in educational applications. The hope of many school districts is that, in the long run, the state can do some of the work for the districts; working on a common data base providing the department with needed statewide information in a single electronic language.

The purpose of this paper to: (1) investigate the need for common data banks in California junior colleges, (2) determine the present development of information systems and data base design, and (3) determine the possible place of generalized file management systems in the development of junior college information systems. To further the writer's understanding of these software packages, forty hours of lecture were attended at the IBM Education Center in a class called "Information Management System."

II. NEED FOR A COMMON DATA BASE IN CALIFORNIA JUNIOR COLLEGES

In 1967, the Danforth Foundation made a grant to sponsor a five-year program for the purpose of initiating, incorporating, and coordinating cooperative projects in Southern California junior colleges. Dr. Sheldon, co-Director of the UCLA Danforth Junior College Program, stated that early in the project it became evident that in order to implement cooperative projects, the same data was required from all participating colleges. Investigation of junior colleges in Southern California indicated that the collection and storing of similar data on their students, faculty, or college business was not a matter of routine (6:preface).

Many colleges fight the movement toward common data gathering because of the fear of losing institutional control and individuality. Yet, if we analyse the amount of required data that is now gathered to answer state and federal questionnaires and reports, the fear of losing individual identity and control becomes, as Dr. Carhart suggests, an academic question (16:9). The pressures of growing enrollments, curricular modernization, and shortage of funds no longer allow educational institutions the luxury of trial and error procedures. Unfortunately, the economics of any one college completing all the programming it needs to automate data for research studies and for fiscal and student personnel procedures is prohibitive.

> The economic facts of life are pressing all of us to search for economical means to solve our problems. In order to cooperate with other colleges, common use of the same data collection systems and definitions is required. . . . The amount of time, money, and energy that can be saved is enormous if we can be assured that completed studies have a common data base. (16:11)

Recently, there have been several projects funded which attempt to experiment with the cooperative sharing of computer resources as well as the cooperative sharing of collected data. An early project described in educational literature is the Suburban School Services Joint Board's "Total Information for Educational Systems." This is a Minnesota $450 thousand project in which educational data will be automatically generated and supplied to the state department of education (13:37). The "Area IX Total Information System" is a $307 thousand project in Scott County, Iowa. The objective of creating a multi-county, multi-district information retrieval project is typical of the several similar projects being funded today (13:38).

The planning committee for the Southern California Junior College Advisory Council recently studied the feasibility of a common data bank and a common data processing system for the California Junior College System. They made the following important recommendations:

1. A system based on the establishment of regional centers formed and operated by a group of cooperating junior colleges. A optimum size for a regional center would permit it to support 5 to 10 colleges (25,000 to 100,000 students); 8 to 12 such centers would be required for the state (6:3-2)

   Advantages:
   - Increased system capabilities such as a larger data base; greater flexibility, reliability, and over all performance.
   - Increased standardization.
   - Lower costs for equipment operation and maintenance.
   - Comparative information available among colleges. (6:3-1)

2. Functional application, sequencing priorities, and schedules for implementation. (6:5-2)

This study also pointed out that some junior colleges would more quickly have the need and the capacity to implement the suggested functional applications. It was the recommendation of this study that in the near future certain junior colleges act as pilot systems (6:5-1). When such a commitment is made at the state level for a pilot system, an important decision will have to be made as to the software used to design, manage, and utilize that data base. The use of generalized file management system software is a possible approach to data base design.

III. AN APPROACH TO DATA BASE DESIGN: MANAGEMENT INFORMATION OR INFORMATION MANAGEMENT?

Existing files have historically been structured for individual storage within core and within secondary storage. Because the structure is for individual applications and because common keys are then the only means of cross referencing two files, there tends to be a great deal of redundancy. Besides the storage costs for the duplication of information in a variety of files, there is the increased possibility of duplicate data items not being updated on a certain file (10:26).

Program maintenance is hampered by the present file structure, in that any change to the file format necessitates a change to the applications program. In addition the programmer must deal with the entire record even though he might be working with only a small segment. This increases the complexity of the program, uses more core, and generally raises programming costs (10:37).

Junior college administrative systems are required to collect certain defined data in order to answer state and federal questionnaires. This being the case, the defined data base itself is quite similar in junior colleges throughout the state. Considerable work has already been done by groups, such as the Southern California Junior College Advisory Council, to determine future research, administration, and educational data needs. Junior colleges in California have, to a varying degree, already structured files containing data for future needs. The problem lies in the fact that the structure and the maintenance of that data base differs widely among these junior colleges in California.

For this reason an immediate objective might be an "information management" system for California junior colleges. Such a system might provide a means of creating and maintaining an information base from present disjointed file structures in an economical and efficient manner.

It is the writer's observation that what most businesses today call a "management information" system, is really a method of achieving economical methods of collecting, transporting, processing, and displaying information (information management).

> This information is not necessarily related to the organization and production of information to support management decision making . . . in fact much of the information being handled by today's automatic systems is valueless for this purpose. (10:41)

Not that all systems experts agree that the requirements of a policy-making function must be included in the development of management information systems.

> The management level planning [represented by the activity of establishing goals and objectives] will probably never be sufficiently systematized to become an integral part of a practical management information system. It may soon become economically feasible, however, to store huge volumes of data in computers and it is this possibility that gives rise to the speculative talk about "selective interrogation of data banks"--since, to be of any practical value, stored data must be retrieved just as information must be intelligently extracted from a reference book. (5:52)

It seems to the writer that the seemingly different concepts of "information management" and "management information" are very closely related and, in fact, are rather dependent on each other. First "information management" must be established in a junior college defined data bank, thus allowing useful "management information" to be developed.

> About 90 percent of our efforts and money go toward getting the data itself. With a [well organized and efficient] data bank we'll be able to use that 90 percent of time to getting answers to important educational questions. (16:15)

Investigation of "information management" systems used in industry suggests some general requirements for such a system to be useful in the California Junior College System.

1. It should permit the establishment of one file with a single area for common information. Information in the . . . [file, such as name and student number], would be recorded once and segments would be added for . . . [personal data, educational achievement, etc.].
2. It should allow certain information to be recorded on disks while sequentially processed or while secondary records are maintained on tape or cards.

> 3. In providing the means to segregate the data base it should also permit guaranteed accuracy of updating by an automatic maintenance feature for all segments of the file.
> 4. It is mandatory that the system permit variable length records so that students with many entries . . . [or academic qualifications, etc.] would consume more file space than new students with few academic qualifications.
> 5. It should permit ease of expanding or reducing the file to add both new entries such as new students and new areas such as additional [academic] qualifications for any given individual. (10:27)

The Information Management System (IMS) announced by IBM is what the writer would call an "information management" system. It is a data management system which segments an individual information system by accessing the data base outside of the application program. Additions can then be made to the file structure without necessitating changes to the hundreds of programs that refer to the data base in general. The processing of a data base organized by such an "information management" system has the additional following advantages:

> 1. Elimination of redundant data on file.
> 2. Reduction of file maintenance.
> 3. Reduction of errors due to updating.
> 4. Consistency of information throughout the files.
> 5. Reduction of data processing costs.
> 6. Reduction of programming maintenance in that the application program is independent of the file organization.
> 7. Less complex program logic, in that most of the I.O. functions are performed by the generalized software system. (10:28)

In addition to software packages that manage data, the junior colleges should also consider the recently developed "generalized management information system" software packages. These generalized systems are variously referred to as "file management systems," "management information systems," and "information languages" (19:24). Essentially they are task-oriented languages rather than computer oriented languages. One writer's opinion was that at least eighty percent of the applications encountered in data processing could be implemented without any formal programming being required if what is referred to in this paper as a "generalized file management system" (GFMS) were used (19:24).

Like a "data management" system, a GFMS also provides the primary functions of file creation, and file maintenance. In addition, however, provision is made for information retrieval and report generation tasks that can be handled by users rather than high level programmers. Many software houses have undertaken the development of GFMS software, urged on by the shortage of qualified programmers, the more complex architecture of third generation machines, and the lengthy implemention periods for relatively straight-forward applications.

Many software executives feel that a GFMS package is a means of circumventing the programming bottleneck faced by most computer users (19:22). The development of these new software packages allow data processing managers to respond to executives who want to obtain data on demand.

Historically this could not be easily or inexpensively done with fixed programs. Even though these new GFMS software packages are expensive to develop, because many customers can use them the cost of development is spread over many users. The price to anyone individual user then becomes quite reasonable (15:31).

These generalized systems are developed with the knowledge of the importance of the man-machine interface. For this reason these new programs are often made easier for unskilled operators to use by causing the system to report back if an incorrect command is encountered. This interface between man and program becomes critical when the GFMS is to be a real time system. It is this writer's opinion that the man-machine interface will be and should be one of the most critical factors in selecting a GFMS software package.

The feasibility study for a common data bank presented by the California Junior College Association recommends the use of general software packages and makes note of the following possible advantages:

1. Applications on older computers can be converted for use on newer models faster, and generally more economically than by reprogramming.
2. Packages are better debugged and more widely field tested when delivered than are most inhouse programs.
3. Documentation is generally superior to that developed by a user's own programmers, thus facilitating conversion and future changes.
4. Package suppliers usually continue to maintain the package against changed requirements.

5. Availability of package suppliers personnel who are generally more versatile and highly skilled than those of the user. (6:3-4)

The next section of this paper investigates the characteristics of a "data management" system and then several "generalized file management systems." Until only recently, this type of software was supplied and supported by the computer manufacturer at no extra charge for the customer. For this reason many customers have experimented with the manufacturer's generalized software, and therefore provided a pool of industrial experience for the uninitiated user. For this reason, the manufacturer's software will be analysed in depth. Advantages as well as disadvantages and shortcomings of the manufacturer's generalized software packages will be noted for California junior colleges.

The number of commercial GFMS software packages competing with the manufacturer's software is growing at an incredible rate and will grow at an even greater rate within the next couple of years. The author will attempt to note the area where competitive commercial systems hope to improve on the manufacturer's generalized software. Based upon what has been done in the past and what is claimed for the future by GFMS software packages, the author will attempt to specify some realistic selection criteria for a junior college system.

IV. GENERALIZED FILE MANAGEMENT SYSTEMS
SUPPORTED BY COMPUTER MANUFACTURERS

Information Management System (IMS)

This system is best described by noting that it manages data more than it provides data to management. IBM describes it as a processing program designed to facilitate the implementation of medium to large common data bases in a multi-application environment. Both online message processing and conventional batch processing are separately or concurrently possible. A suggested strong point of the system is that it permits the evolutionary expansion of data processing applications from a batch only to a high volume information request teleprocessing environment (14:1). The objectives set for the system are the following:

1. To provide data organization methods that are conducive to the creation and maintenance of large common data bases and the multiapplication use of these data bases.
2. To provide the means to permit the user to facilitate development and maintenance of a data base system in the batch processing environment.
3. To provide the user with the ability to extend his data base processing to the teleprocessing or data communication environment.
4. To provide the user with an efficient telecommunication ability for developing a high volume/rapid response online system. (14:4)

IMS/360 is comprised of a data base facility called

Data Language-1. It also contains a data communication facility which includes telecommunications, message scheduling, checkpoints for the restructuring of data from the checkpoints.

The characteristics of Data Language-1 are the following:

1. The ability to create, maintain, and expand the data base.
2. The ability to handle variable length information.
3. Provides the means to make the data and program independent.
4. The ability to handle batch and online processing concurrently or independently. (14:6)

Data Language-1 must be used to define the logical data structure of the data base. This is done external to the applications programs. It is then the responsibility of the applications programmer to provide common source programs linkage so that Data Language-1 sub-programs can be "called" to process input/output requests during execution of the applications programs.

The application programmer must write in one of three programming languages: COBOL, PL1, or ALC. This writer found the data base language DL1 to be quite similar to 360 Operating System Job Control Language. Although the language is essentially a macro language, it definitely is not a language easily mastered by non-programmers. In addition to mastering Data Language-1, the user must also have a knowledge of an applications language such as COBOL to access the data base and retrieve information.

The data base itself is composed of data base records--a data base record being a collection (a variable number) of hierarchically related, fixed length data elements, called segments. A "root" segment is the highest hierarchical segment in the data base record. The writer found the "root" segment to be similar to the FD level indicator in the COBOL programming language. A dependent segment is a data base segment that relies on at least the root segment for its full hierarchical meaning, therefore always being at a lower hierarchical level than the root segment. The writer feels that a dependent segment is quite similar to the 02 level indicator in COBOL. There can be 255 segment types within a data base and fifteen levels of segment hierarchy. Although there is a maximum of 255 segment types or names, a single segment type can occur any number of times and this does not seem to be a major limitation.

IMS/360 allows the user to attach information to a data base and makes it impossible for others to use or access the information. This characteristic is called "sensitivity" to one's own data. The writer finds this to be an adequate method of providing security for a data base.

Minimum system requirements are 128K bytes for batch-only processing and 256K bytes for teleprocessing with IMS. In the later case the system, for all practical purposes, would be dedicated to online IMS/360 applications without the concurrent running of batch-type jobs. In effect then, to use the capabilities of IMS/360 a System/ 360 Processing Unit Model 1 (512K) with Multiplexor Channel, 10 tape units, two 2314 Direct Access Devices would be a very typical system (14:31). The need for such a large machine configuration causes a severe limitation on the usefulness of IMS/360 for the California Junior College System. The writer knows of only one junior college in California that has equipment or plans to have equipment large enough to handle the IMS/360 software package.

If California junior colleges adopt something like regional data processing centers, the availability of an adequate machine configuration for use is more feasible.

The writer's experience with IMS/360 found it to be truely a general purpose system which could be applied to any education organization as well as any company system. This characteristic is not unique to IMS/360. Neither is the fact that applications programs are given independence from the physical organization of the data. However, the method of handling variable length data, providing an adequate security system for data bases, and providing extensive checkpoint and restart facilities is an unusual characteristic for generalized file management systems.

The system designers seem to have been quite concerned about permitting an evolutionary expansion of data processing applications from the batch environment to the teleprocessing environment. Because of this concern much statistical information is automatically collected to assist the user in evaluation of performance and changing communication requirements.

This system became available from IBM in the summer of 1969 at a rental cost of $400 a month. This system could possibly be utilized by a regional group of junior college with a large enough machine configuration or at a central state facility.

Generalized Information System (GIS)

As a system IBM's GIS provides facilities for defining, maintaining, and retrieving data from user files under the direction of the using installation (11:1). GIS is purported to be designed for managers who need specific information as well as answers to "what-if" type of questions. Under GIS the computer is to be able to draw upon interrelated data from different information files in order to develop answers (4:23).

GIS functions are initiated and controlled by database-oriented macro instructions that constitute the problem language. Although this is purported to be a near-English language, many users feel that the ease of management useage only takes place if that management has had actual programming experience (17:43). Nevertheless, the GIS language is made up of descriptive, procedural, and control commands which can work on files created by ALC, PL1, or COBOL programs. GIS "descriptive" statements explain the design of each logical file in the data base and are processed to form a permanently stored control table that contains all logical physical characteristics of a file. A file is thus described only once, and its use in any subsequent task is accomplished by means of symbolic identifiers. Procedural statements indicate the action to be taken and the conditions under which they are to be performed. A coherent set of procedural statements constitute a task specification. Three types of task specifications are possible:

1. File creation--to establish a file for the first time.
2. File update--to enter new data or delete or change data already in a file, and
3. File inquiry--to locate, organize, summarize, and present data. (2:97)

Statements in the control catagory establish operating controls that may modify or regulate the way in which other tasks and processing jobs are performed. Control statements may cause the following:

1. A task specification to be stored in its original or executable form for later or repeated use.
2. A saved task to be executed at a specified time or during a specified interval.
3. A sequence of displays and computer actions to be stored for recall as part of a structured man/machine conversation from remote terminal

> devices, or as elements of similar cue-and-response procedures. (2:98)

In addition to standard Boolean operators, GIS permits four other conditional operators not commonly found in generalized file systems: BETWEEN, SCAN, MASK, and CHANGE. The operator BETWEEN determines whether a field value is between the two values given. The operator SCAN tests a field for a particular pattern of characters. The operator MASK tests a field for a match on certain characters while at the same time allowing other characters to take on any value. Finally, the CHANGE operator allows a test for a change in contents of a field from one record to the next.

The GIS record structure is made up of logical records which may be stored as a single physical record or as a physically split record. In the single physical record's case, the files are organized either as sequential or indexed sequential. The FILE KEY provides the field upon which an order is imposed. In the split-record situation, the master segment is ordered according to the KEY; the trailer segments have no order and are instead bound to the master segment by links and chains. A link provides the means of finding the master segment from a trailer segment. Chains link master to trailer, trailer to trailer and eventually the last of a sequence of trailers back to the master.

GIS's Data Description Tables (DDT) are permanent tables which relate symbolic file, field, and segment names to physical characteristics in the system. The format and content of the DDT as well as the characteristics at record and sub-record level according to core memory available is presented below.

	64K	128K	256K
Record Size	2K	4K	8K
Number of Cross Indexes	5	10	20
Segments	10	30	100
Number of Files Processible	2 (5)	3 (6)	4-6 (12)
Fields	100	250	500

(17:36)

Many system experts consider GIS to be a comprehensive generalized system of significant utility. However, the usefulness of the language has been questioned because of slippages over the last three years as far as full implementation of the system. Poor performance, considering the speed of the equipment, due to its interpretive mode of execution has been a major drawback (19:26).

As far as implementation of GIS on a regional junior college system, the realistic minimum machine configuration size of 128K again might be economically restricting. This writer will look with great interest into GIS (Basic) which is supposed to require a smaller machine configuration size of 65K. The writer found no company presently fully implementing GIS (Basic).

Information Oriented Language (INFOL)

INFOL is a tape-oriented system developed by Control Data Corporation for its 3600 and 3800 series of computers. Implemented in assembly language, it runs under the SCOPE operating system.

The type of user language is free form using field numbers for data identification. Both fixed and variable record formats are accommodated. The operating environments is batch-only with no support for online processing. INFOL is a tape system, thereby emphasizing sequential processing. The minimum machine configuration has a reasonable core size of 65K with five tape drives. Unfortunately, CDC has indicated that no future releases of the system will be made and the existing version will not be supported (19:26).

The characteristic capabilities of INFOL are similar to those of other task-oriented systems and are presented below:

1. The system can be used by persons who have little or no knowledge of programming techniques.
2. Conditions on fields in records include the relatioral criteria, "equal," "greater than," "less than," etc.
3. Output may be defined by the user or left to the system.
4. Validation criteria may be imposed upon input.
5. File revision, necessary after extensive updating (with additions and deletions especially), can be handled with relative ease.
6. Simple computations, including averaging, can be effected through simple system commands.

7. Extensive diagnostic messages (about 250) prevent a variety of miscues. (17:13)

One of the unique aspects of INFOL is its approach to formatting. INFOL stores information in a free-field fashion, therefore, needing no specification as to type and number of characters in a field. INFOL, therefore, can process both fixed and variable length fields because they are handled precisely the same.

The INFOL record itself is not hierarchical in form. There is no restriction on the type of file that can be handled since it is possible to associate fields and nest the association. Because the data must be reformated before the data base can be accessed, INFOL might be more closely compared to IMS/360 than to GIS. Because of the need for reformatting (called ESTABLISHMENT by CDC), some system experts disqualify INFOL as a "Generalized" file management system on the same grounds as they disqualify IMS/360. However, this writer finds it important to note that INFOL qualifies quite nicely as a generalized file management system when judged on ease of customer usage; especially when compared to IMS/360.

The writer finds INFOL to be a system of significant value and recommends that a regional junior college system considers its use if CDC equipment is involved. Unfortunately INFOL, like IMS/360 and GIS, is not readily adaptable to another manufacture's machines. Certainly this factor is a big disadvantage of manufacturer supported generalized file management software packages and a main advantage of some commercial systems. Once designing a junior college Management Information System around a system like IMS/360 or INFOL, change to another computer manufacturer presently becomes very costly and time consumming.

INFOL, GIS, and IMS/360 were the three manufacturer systems studied in depth, analysing users manuals or attending the manufacturer's school. Three other systems are mentioned in the literature and are now briefly discussed.

Business EDP Systems Technique (BEST)

BEST is an integrated system of fifty-five data processing functions which were designed to circumvent the programming of at least 50 percent of the applications encountered by users of NCR's 315 computer. The type of language is tabular, the record formats are fixed, and the processing access method is sequential (19:26).

BEST operates on a relatively small system (NCR 315 and 20,000 character positions) with five tape drives. Like previously discussed manufacturer software, BEST is written in a language unique to the manufacturer (NEAT). This of course makes BEST very difficult to adopt for another system even if the manufacturer would release the software (19:27).

BEST has been praised for its ease of customer usage. Although definitely a modest system, it has been able to surpass the goal of handling 50 percent of a company's typical applications. On the minus side, users report that the compiling of a BEST program is slow. This is explained by the fact that BEST generates a NEAT assembly language program that must also then be assenbled (19:27).

Integrated Data Store (IDS)

IDS is one of the first attempts by a manufacturer to develop a supported generalized program to organize, store, maintain, and retrieve information. IDS is not an information storage and retrieval system of and by itself, but is instead used with COBOL to provide logical record-processing capabilities. IDS resembles IMS/360 in that it is not an inexperienced-user type of language; instead requiring a programmer's knowledge of IDS and COBOL (20:3-275).

IDS language provides data description elements which declare the existence of master detail relationships between records. These relationships are implemented through chain link processing techniques which create circular or ring structures quite like IBM's GIS package. The master record links to the first detail record; the first detail record links to the second detail record; and so on until the last detail record in the chain, which links back to the master record. Records may optionally contain the address of the prior record and/or master record (20:3-231).

IDS operates in either a batch or direct access mode and has the ability to store and retrieve records of any length within a mass memory. The records themself are defined by COBOL data descriptions to contain a specified number or fixed size fields and to participate in a specified number of chains (20:3-234).

The writer finds IDS to be similar to IMS/360 but tied more closely to the compiler language. It seems that IBM borrowed several concepts from this early GE system and then increased the flexibility of their own system by separating the data division entirely from the compiler language. For a COBOL shop using GE equipment IDS is a worthwhile system and should be considered.

MANAGE

MANAGE is a Scientific Data System developed generalized file management system. It has been successfully used in wide variety of applications since early 1967. MANAGE has particularly been praised as a flexible general purpose system which can be implemented for small computer configurations. The core requirements for the SDS 900

Series is 8K (24 Bit words) or 16K for the SDS 9300 Series. Three tape drives, card reader, and line printer are also required (19:26).

The language used by MANAGE is the SDS BUSINESS LANGUAGE which is tabular in form. SDS claims that this language is a machine independent assembly language; a distinct advantage over the previously mentioned systems (19:26).

V. CONCLUSIONS AND RECOMMENDATIONS

One author writing about generalized file management systems made the following very true statement:

> As a general rule (and not too surprising) it is safe to say that what is on the drawing boards is somewhat superior to what has been developed. (17:1)

The writer feels that the above statement represents the crux of the problem regarding the use of generalized file management systems. Many of the manufacturer systems require a large machine configuration which the individual junior college just cannot afford, especially if it is a small junior college. This is one of the reasons that the concept of a common central data bank and regional computer centers makes such good sense. Yet, it is important to note that critics of manufacturer supported software claim a lack of flexibility and/or inefficient performance of some generalized file management systems (19:22). Plans for more efficient software systems as well as for smaller software systems are on the drawing boards of both computer manufacturers and commercial software houses; and offer a potential future solution.

The writer finds it quite disturbing that many systems tie the user into that manufacturer's system or at least into the same family of systems. This of course is disturbing but not surprising. Because the computer at a junior college is so widely used for the two year computer vocational program, many junior college teaching staffs perfer to offer the student training on IBM 360 equipment. Because the same equipment is often also used for administrative functions, many administrative data processing departments rely on software supported by IBM. The writer suspects that IBM's GIS/BASIC will be further refined to the point where it becomes more efficient as well as feasible for a small 360/30. However, an administrative system tied into GIS/BASIC will be difficult to redesign when there is a change to a new manufacturer or to a higher family of computers.

Although IMS/360 is available only for very large 360's, the writer suspects that a type of "IMS/BASIC" will be developed for the type of smaller system that might be found in an individual junior college or regional junior college computer systems. The concept of managing a data base outside of the application program seems to be a sound idea. IMS/360 or any data management system, would hopefully provide sufficient flexibility to move across manufacturer or computer family lines.

The writer is quite hopefull that software houses will fill in the gap left by manufacturer supplied GFMS packages for small computer users. These systems will most likely be both more flexible and more efficient; so as to create a profit motive over a wide range of customers. Until there are a greater number of generalized file management system alternatives, what steps can a junior college system take to utilize present software packages in the creation and use of data base systems?

For junior colleges to band together and create a meaningful common data base, coordination at the state level will be required to point out which hardware systems will be compatible as well as which software systems will be compatible. Companies adopting generalized file management systems have suggested that previous users are often the best source of information on limitations, errors, and other problems associated with using a software package (3:20). It has also been found that an organization should keep the software as near the vendors "off-the-shelf" version as possible to take advantage of the maintenance that the vendor provides as well as to improve the chances of getting proper assistance (3:21). Discussions with present and past users of generalized file management system are a source of information that the individual junior college should utilize. The emphasis is made regarding the involvement of the individual junior college, because the writer feels that the support for state level coordination must first be desired by the individual junior colleges.

Prior to the selection of an appropriate software system the junior college will have to be involved in the large task of designing and redesigning an appropriate common data base. It is hard enough to collect and save data which has been produced for some defined purpose, but what about data needed for future unpredictable purposes? Considerable additional study still must be done to determine what should be required in a common data base for a California Junior College System. Even after this is done a considerable selling job will have to be undertaken to acquire sufficient support and sufficient funds. Hayes in a recent article points out the need for justification:

> . . . whereas the original recording of the data had clearly defined value, related to the purpose of its generation, the utility for other, unknown uses is not all clear. . . . Since this [collection of data] is a "capital investment," it must be amortized, in some sense, over these future, unknown uses made of it. As a result the costs of an information retrieval system are difficult to justify. (7:26)

Even before the need for justification of a data base system, the problem of junior college administrator's reluctance toward a data base system will have to be overcome. This will be true at both the state and individual junior college level; and is certainly not unique to the junior college environment. Corporations who have undertaken a program of building a data base to enable an information system to function, have found the following factors to be a source of confusion and lack of understanding.

1. Associates in other companies either praising or condemning the results of similar programs

undertaken by their companies. This contradiction becomes more confusing when the reasons given for success or failure differ from person to person.

2. Internal line and staff management personnel stating the absolute necessity for, or violent objection to such an undertaking with each endorsement or objection supported with valid proof.
3. Presentations by experts in the information processing field pointing out the simplicity or extreme complexity of such an undertaking.
4. The sudden appearance of excessive costs, during or following installation, that were not considered at the time of the original estimates.
5. Underlying fear of personnel that installation of a system will uncover weak spots in the organization that could lead to changes affecting them. (1:10)

Finally, it is the writer's opinion that the junior colleges can very profitably determine if their own environment is such that the design, installation, and continued growth of an information system is probable. Many systems experts feel that it is best to delay the start of company involvement in a complex automated information system until the following items are in place:

1. Adequate corporate discipline so that common interfunctional procedures can be implemented.
2. Documentation of potential savings anticipated from installation.
3. A relatively stable management, especially at the policy-making level.
4. Management that is willing to commit its own time and interest to understand the various plans, techniques and equipment associated with the proposed system. This understanding should be in sufficient detail to enable intelligent monitoring of the costs and progress of the system.
5. The willingness of management to start acquiring and training a core of experienced systems personnel.
6. The presence within the organization of

> operational personnel who are knowledgeable in depth concerning the information requirements, methods, procedures and techniques within the functions they are associated with. (1:12)

When the junior college has satisfied these reasonable prerequisites related to people, the largest step toward a common data bank has been taken. Once this commitment has been made to install a common junior college data base system, differing means of entering that data base must be measured against some criteria. The writer's investigations of computer-manufacturer-supplied Generalized File Management Systems suggests several broad catagories of criteria: file independence, low problem definition, strong user orientation, and consistency of file processing. All computer manufacturers and all commercial software houses claim to offer a GFMS that meets at least these criteria to some degree. The problem is to determine that degree.

The disadvantages connected with generalized systems supplied by the manufacturer are often rationalized away by an absence of a purchase price or perhaps a very low purchase or lease price. This is often the claim of the enthusiastic salesman from a commercial software house. The truth is that there are additional implementation costs to be considered for all generalized file management systems whatever the source:

1. Training costs

2. Manpower time
3. Machine resources
4. Costs of distributing and maintaining program documentation.

It seems apparent, however, that what a commercial software house might offer is more efficient use of machine resources and a more efficient use of user time in return for a purchase or lease price. Certain factors are noticeably variable among many of the generalized file management systems:

1. Ease of integration with other software systems.
2. The current status of the software package (e.g. testing or production stage).
3. Modification constraints.
4. Number of existing installations.

It seems to this writer that two of these factors are especially vital criteria when considering use in a junior college system. Because of a desire for a degree of autonomy among individual junior colleges or among regional junior college systems, it is vital that a GFMS be compatible with a great variety of software systems and modification constraints are flexible enough to adapt to various administrative systems.

Many generalized file management systems vary as to file and record definition. Because various file structures must be accommodated in a junior college system, the following are recommended as criteria in selection of a GFMS.

1. The ability to accommodate hierarchical relationships in raw data.
2. The ability to overlay descriptions on existing files.
3. The ability to accommodate fixed, variable, and undefined record formats.

Most generalized file management systems claim to be designed for the convenience of non-programmers. In comparing specific systems the following criteria should be considered to enable easy use by non-programmers.

1. Shorthand default options.
2. The facility for synonym searching.
3. The cataloging of requests for future use.
4. The ability to print successively a number of reports in one pass through the file.
5. The existance of extensive diagnostics.

Many marketable generalized file management systems place a heavy emphasis on their report-generating capabilities. There are certain criteria that can be used in judging this capability:

1. Flexibility in choice of output medium.
2. Flexibility in format.
3. Automatic security messages when data defined as confidential is retrieved.

The writer has earlier mentioned processing efficiency as one of the strongest points of a marketable GFMS for junior colleges. What then are the criteria for measuring this efficiency of processing?

1. The ability to effectively operate in a relatively small system environment.
2. The ability to include task programs written in a procedural language.
3. The ability to constantly check for data validity.
4. Data integrity.
5. The ability to retrieve data from multiple-files during a single run.

Historically, customer support has proven to be an important factor in effectively installing a GFMS. The following criteria should be used as indicators of sufficient customer support.

1. Sufficient documentation available on the system.
2. Training requirements for users adequately planned for by the vendor.
3. Service response on problems meets the user's requirements.
4. Maintenance support is included in the price of the package.

In summary, the design of a common junior college data base is first required before a Generalized File Management System can be considered. A central state data base system requires a user oriented method of accessing the data base as well as an efficient method of managing the data. As this writer sees the problem; the computer manufacturer makes available a system to efficiently manage the data base, but does not offer a system tailored to the non-technical user. The solution might be to efficiently manage the data base with provided software while evaluating commercial Generalized File Management Systems.

The writer has indicated what he considers to be important criteria in evaluating processing efficiency, factors of file and record definitions, convenience of use for non-programmers, report generating capabilities, and customer support. When the common data base system has been designed, representatives from the individual junior colleges or from regional computer centers should be polled to include other possible criteria. The expense of evaluating the growing number of Generalized File Management Systems could be absorbed once by a central state agency. The purchase cost of such a system should be absorbed at the state level and the generalized software package made available to members of the junior college network. This idea of evaluating Generalized File Management Systems can be undertaken by regional data processing centers but without the economy of scale possible at the state level. It is quite possible that more than one Generalized File Management System would be required to meet certain criteria. Only at the state level can the purchase or lease of more than one of these generalized systems be an economic reality.

The realistic sequence of activities must then be the creation of a common junior college data base, the creation of a central information network, criteria specified for a Generalized File Management System; and finally, availability of this generalized system to access, retrieve, and update the data base itself.

BIBLIOGRAPHY

1. Barnett, Joseph I. "How to Install a Management Information and Control System." Systems and Procedures Journal, XVII, No. 4 (September-October, 1966), 10-14.

2. Bryant, J., II and Parlan Simple, Jr. "GIS and File Management." Proceedings--A.C.M. National Meeting, 1966, pp. 96-107.

3. Burke, Duane and Robert Gillespie. "Purchasing Package Software--A Customer's Point of View." Computers and Automation, February, 1969, pp. 19-22.

4. "Data For Management." Data Processor, XI, No. 5 (November, 1968), 23-24.

5. Donkin, Robert G. "Will the Real MIS Stand Up?" Business Automation, XVI, No. 5 (May, 1969), 50-53.

6. Feasibility Study Common Data Bank and Common Data Processing System for California Junior College Association. Research and Development Committee California Junior College Association, October, 1968.

7. Hayes, Robert M. "Information Retrieval: An Introduction." Datamation, XIV, No. 3 (March, 1968), 22-26.

8. Head, Robert V. "Information Systems: The Changing Scene." Journal of Systems Management, II, No. 4 (April, 1969), 40-41.

9. Head Robert V. "Structuring the Data Base for Management Information Systems." Journal of Systems Management, XX, No. 1 (January, 1969), 9-11.

10. Gallery, Thomas A. "An Approach to Data Base Design." Journal of Systems Management, XX, No. 2 (February, 1969), 26-28.

11. *Generalized Information System (BASIC) Application Description*. IBM Doc. No. 20-0521-0, 1968.

12. *Generalized Information System, Application Description*. IBM Doc. No. 20-0574-0, 1968.

13. Grant, C. B. S. "Federal Grants Encourage Regional Centers Total Systems." *Data Processing Magazine*, X, No. 7 (July, 1968), 36-37.

14. *Information Management System/360 for the IBM System 360, Application Description*. IBM Doc. No. H-20-0524-1, 1968.

15. Kircher, Paul. "Breakthrough in Management Information Systems." *Data Management*, February, 1969, pp. 28-31.

16. Nemefee, Audrey, ed. Development and Use of a Common Data Bank. Committee on Research and Development, California Junior College Association, March, 1968.

17. Reilly, Kevin D. "Evaluation of Generalized File Management Systems." *Final Report on Mechanized Information Services in the University Library. Part 6.* Los Angeles, California: Institute of Library Research, University of California, 1967.

18. Report - UCLA Danforth Junior College Program. Graduate School of Education, University of California, Los Angeles, February, 1969.

19. Sundeen, Donald H. "General Purpose Software." *Datamation*, January, 1968.

20. Bachman, Charles W. "Integrated Data Store." *Proceedings of the Second Symposium on Computer Centered Data Base Systems*, December, 1968, pp. 231-275.